PB83-131243

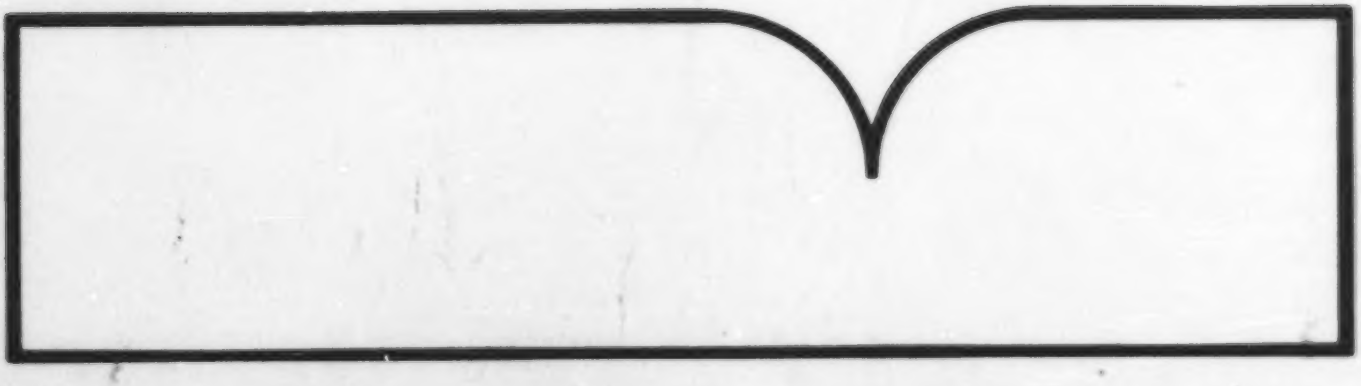
Robust Estimates of the ED50

Montana State Univ.
Bozeman

Prepared for

Environmental Research Lab.-Duluth, MN

Jun 79



U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completi

1. REPORT NO. EPA-600/J-79-145	2. Journal Article
3. PB83-131243	
4. TITLE AND SUBTITLE Robust Estimates of the ED50	5. REPORT DATE 1979 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Hamilton, M. A.	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematical Sciences Montana State University Bozeman, Montana 59717	10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO. R803950
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency Environmental Research Laboratory-Duluth 6201 Congdon Boulevard Duluth, Minnesota 55804	13. TYPE OF REPORT AND PERIOD COVERED 14. SPONSORING AGENCY CODE EPA/600/03
15. SUPPLEMENTARY NOTES J. Am. Statist. Assoc. 74(366): 344-354 (1979)	

16. ABSTRACT

To explore the possibility that robust estimators of location can be adapted to the problem of estimating the ED50 in binary-response bioassay, 10 estimators are compared. A Monte Carlo study is conducted to determine the mean squared errors (MSE) of the estimators. Taking into account the MSE for a variety of symmetric tolerance distributions, the sensitivity of the estimator to an anomalous response, and the possibility that the estimator is incalculable, the trimmed Spearman-Karber estimator is especially reliable. Adjusting the response proportions before calculating the ED50 significantly reduces the MSE of the minimum logit chi-squared estimator.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS *(This Report)* UNCLASSIFIED 20. SECURITY CLASS *(This page)* UNCLASSIFIED	21. NO. OF PAGES 13 22. PRICE

EPA-600/J-79-145
JOURNAL ARTICLE

Journal of the American Statistical Association
June 1979, Volume 74, Number 366
Theory and Methods Section
Pages 344-354

Robust Estimates of the ED50

MARTIN A. HAMILTON*

Robust Estimates of the ED50

MARTIN A. HAMILTON*

To explore the possibility that robust estimators of location can be adapted to the problem of estimating the ED50 in binary-response bioassay, 10 estimators are compared. A Monte Carlo study is conducted to determine the mean squared errors (MSE) of the estimators. Taking into account the MSE for a variety of symmetric tolerance distributions, the sensitivity of the estimator to an anomalous response, and the possibility that the estimator is incalculable, the trimmed Spearman-Karber estimator is especially reliable. Adjusting the response proportions before calculating the ED50 significantly reduces the MSE of the minimum logit chi-squared estimator.

KEY WORDS: Binary-response Bioassay; Spearman-Karber; trimmed Spearman-Karber; Maximum likelihood; Minimum logit chi-squared; *M* estimates.

1. INTRODUCTION

Suppose that many sets of binary-response bioassay data are to be analyzed on a routine basis and that the only parameter of interest is the median effective dose (ED50). The experimenter may ultimately plot or otherwise compare the estimated ED50's and request that all estimates of the ED50's be based on a single statistical method. This estimation problem arises, for example, when one is routinely monitoring levels of chemicals that are toxic to fish by using the recommended bioassay procedures (Committee on Methods for Toxicity Tests With Aquatic Organisms 1975). In such experiments the doses of the toxicant are widely spaced relative to the underlying dose-response curve.

The appropriate statistical method for these routine analyses must have the following characteristics: (a) be reasonably accurate and precise; (b) be relatively unaffected by a small fraction of anomalous responses; and (c) be calculable for all, except the obviously pathological, experiment outcomes. The popular methods are deficient in one or more of these characteristics. Some deficiencies are described in Section 3 (also see Hamilton, Russo, and Thurston 1977, 1978; Stephan 1977).

The purposes of this article are to describe the inadequacies of some popular methods and to discuss some new nonparametric methods for estimating the ED50. These new methods are checked for the three required characteristics and are compared with the popular methods. The nonparametric methods are simple adaptations of procedures proposed in the literature as robust estimates of the location parameter when sampling from a symmetric frequency distribution (Andrews et al. 1972).

Section 2 of this article contains notation and descriptions of all estimators to be studied. Section 3 presents hypothetical data to illustrate some deficiencies of the methods and contains some sensitivity analyses that tell how seriously each estimator is affected by a single anomalous response. A Monte Carlo study of the mean squared errors (MSE's) of the estimators is described in Section 4. Recommendations and discussion make up Section 5.

2. DESCRIPTION OF ESTIMATION METHODS

2.1 Preliminaries

The bioassay experiment is conducted by giving selected doses of a drug to subjects and observing whether each subject responds. The experiment yields the following information:

1. $x_1, x_2, \ldots, x_k$, which are the k doses of the drug used and which are arranged in increasing order so that $x_1 < x_2 < \ldots < x_k$;
2. $n_1, n_2, \ldots, n_k$, which are the numbers of subjects given doses $x_1, \ldots, x_k$, respectively;
3. $r_1, r_2, \ldots, r_k$, which are the numbers of subjects that respond to doses $x_1, \ldots, x_k$, respectively.

The observed proportion of responses to dose x_i is denoted by $s_i = r_i/n_i$, $i = 1, \ldots, k$.

Suppose that the subjects used in the experiment were drawn at random from some large population of subjects. Assume that each subject in the population has a unique upper limit of dose tolerance, which is that dose just sufficient to cause the subject to respond; at any dose less than the tolerance dose the subject would not respond. The frequency distribution of tolerance doses for all subjects in the population is called the tolerance distribution.

Denote the median of the tolerance distribution by μ, a parameter commonly called the median effective dose or simply the ED50. All the statistical methods discussed in this article require that the dosage scale is chosen so that the population tolerance distribution is symmetric

* Martin A. Hamilton is Associate Professor, Department of Mathematics, Montana State University, Bozeman, MT 59717. He completed this research during 1977–1978 while at the National Institute of Environmental Health Sciences at Research Triangle Park, North Carolina. This work was funded in part by the U.S. Environmental Protection Agency, Environmental Research Laboratory–Duluth, Research Grant No. R803950. The author benefited from discussions with Drs. R. Lund, M. Layard, and D. Hoel, and from comments by anonymous referees.

around μ. Thus, μ is also the mean of the tolerance distribution whenever the mean exists.

The true response proportion $P(x)$ is that (unknown) proportion of the population of subjects that would respond if given dose level x. Notice that the response curve $P(x)$ is the cumulative relative frequency curve (i.e., the distribution function) for the tolerance distribution; $P(\mu) = \frac{1}{2}$, and for every $\delta > 0$, $P(\mu - \delta) = 1 - P(\mu + \delta)$.

The observed proportion of responses, s_i, is an estimate of $P(x_i)$. Under the assumptions just presented, the probability that the number of responses at dose x_i is r_i is given by the binomial probability

$$\text{pr}(r_i) = (n_i!/r_i!(n - r_i)!)[P(x_i)]^{r_i}[1 - P(x_i)]^{n_i - r_i} \quad (2.1)$$

and the joint probability of the responses $r_1, \ldots, r_k$ is given by

$$\text{pr}(r_1, \ldots, r_k) = \prod_{i=1}^{k} \text{pr}(r_i) \ . \quad (2.2)$$

Equations (2.1) and (2.2) were used in the generation of data for the Monte Carlo study of Section 4.

2.2 Description of Estimators

Ten estimators are to be discussed in this article. All but two of the estimators require an initial adjustment of the observed proportions $s_1, s_2, \ldots, s_k$ whenever the ordering $s_1 \leq s_2 \leq \ldots \leq s_k$ is violated. The adjustment amounts to finding the maximum-likelihood estimators (MLE's) of $P(x_1), \ldots, P(x_k)$ as described in Barlow et al. (1972, Example 2.8). Define adjusted response proportions $p_1, p_2, \ldots, p_k$ by combining the number of responses (r_i's) and the number of subjects (n_i's) of any adjacent s_i's that are not in the proper nondecreasing order to give a new averaged estimate, p, at the two adjacent doses. Thus, if $s_i \leq s_{i+1}$, then $p_i = s_i$ and $p_{i+1} = s_{i+1}$, but if $s_i > s_{i+1}$, then $p_i = p_{i+1} = (r_i + r_{i+1})/(n_i + n_{i+1})$. This process is continued until the p_i's are in a monotone nondecreasing order. The general formula (Barlow et al. 1972, p. 19) is

$$p_i = \max_{1 \leq u \leq i} \min_{i \leq v \leq k} \left[\sum_{j=u}^{v} r_j / \sum_{j=u}^{v} n_j \right] .$$

If one plots the (x_i, p_i) points and connects successive points with straight lines, a piecewise linear function is formed. This polygonal segment, denoted by $\hat{P}(x)$, is an estimate of $P(x)$ for $x_1 \leq x \leq x_k$. All seven nonparametric estimators in this article are functions of $\hat{P}(x)$.

2.2.1 MLE.

Suppose $P(x)$ follows the logistic model

$$P(x) = [1 + \exp\{-(\alpha + \beta x)\}]^{-1} \ .$$

Then one finds the following normal equations, in which a_1 and b_1 are the MLE's for α and β:

$$\sum_{1}^{k} n_i[1 + \exp\{-(a_1 + b_1 x_i)\}]^{-1} = \sum_{1}^{k} r_i \quad (2.3)$$

and

$$\sum_{1}^{k} n_i x_i[1 + \exp\{-(a_1 + b_1 x_i)\}]^{-1} = \sum_{1}^{k} r_i x_i \ .$$

The MLE of μ is $\hat{\mu} = -a_1/b_1$.

The iterative procedure described by Ashton (1972, Method 2, p. 27) was used in all calculations of this article to find solutions to (2.3). The initial guess of a_1, b_1 was found by Equation (6.28) of Cox (1970). For these calculations, MLE was said to be calculable if and only if convergence occurred within 30 iterations. The iterations were said to converge whenever the relative changes of both a_1 and b_1 were less than .01 percent. As is typical with iterative MLE algorithms, the calculated MLE may not be identically the MLE.

Two situations in which MLE is not calculable by this iterative procedure deserve special mention. If $s_1 = \ldots = s_{j-1} = 0$ and $s_j = \ldots = s_k = 1$, then the MLE of μ is not unique and the MLE of β is infinity (i.e., the MLE of the variance of the logistic tolerance distribution is zero). If $s_j \neq 0$ or 1, $s_1 = \ldots = s_{j-1} = 0$, and $s_{j+1} = \ldots = s_k = 1$, then the MLE of μ is x_j and the MLE of β is infinity. In either of these cases the iterative procedure will not converge. The MLE program could be altered to detect those sets and find an exact MLE without going through the iterative procedure. In these cases, however, because MLE declares the tolerance distribution to be degenerate, no estimated standard error of $\hat{\mu}$ can be calculated.

2.2.2 *Minimum Logit Chi-squared Estimator* (MLCS).

The logit transformation of $P(x)$ is defined by $\text{logit}(P(x)) = \ln[P(x)/(1 - P(x))]$. If $P(x)$ follows the logistic model, then $\text{logit}(P(x)) = \alpha + \beta x$. Let ℓ_i be an estimate of $\text{logit}(P(x))$ based on the bioassay data and use as a measure of distance between the true and observed logits, the logit chi-squared measure (Ashton 1972) $\chi^2 = \sum_1^k W_i(\ell_i - \alpha - \beta x_i)^2$, where

$$W_i = n_i P(x_i)(1 - P(x_i))$$

is the reciprocal of the asymptotic normal variance of ℓ_i. Then the MLCS's of α and β, denoted by a_2 and b_2, are the values that minimize χ^2. The calculations of this article used the empirical logit $\ell_i = \ln\{(s_i + 1/2n_i)/(1 - s_i + 1/2n_i)\}$ and used, in place of W_i, the empirical weight

$$w_i = n_i\{[1/(s_i + 1/2n_i)] + [1/(1 - s_i + 1/2n_i)]\}^{-1} \ .$$

The weight w_i was set to zero if either $s_1 = s_2 = \ldots = s_{i+1} = .0$ or if $s_{i-1} = s_i = \ldots = s_k = 1.0$ (Ashton 1972, p. 39).

The MLCS estimates are identical to the intercept and slope of a weighted least squares regression line, where ℓ_i is the dependent variable, x_i is the independent variable, and w_i is the weight. The equations for a_2 and b_2 are

$$b_2 = \sum_{1}^{k} w_i(\ell_i - \bar{\ell})x_i / \sum_{1}^{k} w_i(x_i - \bar{x})^2$$

and

$$a_2 = \bar{\ell} - b_2\bar{x} \ , \tag{2.4}$$

where

$$\bar{x} = \sum_1^k w_i x_i / \sum_1^k w_i \quad \text{and} \quad \bar{\ell} = \sum_1^k w_i \ell_i / \sum_1^k w_i \ .$$

The associated ED50 estimate is $\hat{\mu} = -a_2/b_2$. Calculation of MLCS is straightforward unless b_2 nearly equals zero, a situation that is highly unlikely if the doses are so widely spread that $P(x_1) \doteq 0$ and $P(x_k) \doteq 1$.

2.2.3 MLCS *After Monotone Adjustment of the Response Proportions* (MLCSM).

The estimator is $\hat{\mu} = -a_3/b_3$, where a_3, b_3 are given by (2.4) except that in the definitions of the logit, ℓ_i, and the weight, w_i, the proportion s_i is replaced by p_i. As with MLCS, MLCSM is nearly always calculable when the doses are widely spaced.

2.2.4 Spearman-Karber Estimator (SK).

When the mean of the population tolerance distribution exists,

$$\mu = \int_{-\infty}^{\infty} x dP(x) \ .$$

The SK is

$$\hat{\mu} = \int_{-\infty}^{\infty} x d\hat{P}(x) \ .$$

If $p_1 = .0$ and $p_k = 1.0$, $\hat{\mu}$ can be written

$$\hat{\mu} = \sum_1^{k-1} (p_{i+1} - p_i)(x_{i+1} + x_i)/2 \ .$$

The conventional rule for calculating SK if $p_1 > .0$ or $p_k < 1.0$ is described and criticized in Finney (1964). For purposes of this article, the conventional rule was not followed; instead, the SK estimator was said to be incalculable if $p_1 > .0$ or $p_k < 1.0$.

2.2.5 Trimmed SK *Estimators* (*5%*SK, *10%*SK, *20%*SK, *and 50%*SK).

These four estimators are related to the SK estimator in the same way that the trimmed mean is related to the sample mean (Andrews et al. 1972, p. 31). Let A be a chosen constant, $.0 \le A \le .5$. Let

$$L(A) = \max\{i: p_i \le A, i = 1, \ldots, k\}$$

and

$$U(A) = \min\{i: p_i \ge 1 - A, i = 1, \ldots, k\} \ .$$

Let

$$\hat{P}^{-1}(A) = x_{L(A)} + (x_{L(A)+1} - x_{L(A)})(A - p_{L(A)})/(p_{L(A)+1} - p_{L(A)}) \ ;$$

$$\hat{P}^{-1}(1-A) = x_{U(A)-1} + (x_{U(A)} - x_{U(A)-1})(1 - A - p_{U(A)-1})/(p_{U(A)} - p_{U(A)-1}) \ .$$

Then the $100A\%$ trimmed SK estimator of μ is

$$\begin{aligned}\hat{\mu}(A) &= \int_{\hat{P}^{-1}(A)}^{\hat{P}^{-1}(1-A)} x d\hat{P}(x)/(1 - 2A) \\ &= [\hat{P}^{-1}(A) + x_{L(A)+1})(p_{L(A)+1} - A)/2 \\ &\quad + \sum_{L(A)+1}^{U(A)-2} (x_i + x_{i+1})(p_{i+1} - p_i)/2 \\ &\quad + (\hat{P}^{-1}(1 - A) + x_{U(A)-1}) \\ &\quad \cdot (1 - A - p_{U(A)-1})/2]/(1 - 2A) \ . \end{aligned} \tag{2.5}$$

For the purposes of this article, this estimator is said to be calculable only if $p_1 \le A$ and $p_k \ge 1 - A$.

The estimators 5%SK, 10%SK, 20%SK, and 50%SK are given by (2.5) with $A = .05$, $.10$, $.20$, and $.50$, respectively.

One could identify SK as a 0% trimmed SK estimator. The calculations of 50%SK amount to linear interpolation between the two central dose-response points for which the p's enclose .50.

2.2.6 The One-Step Huber M *Estimator With Robust Scale and Constant = 1.5* (P15).

This estimator, which has been found to be particularly useful in one-sample problems, is described in Andrews et al. (1972). For purposes of this article a pseudosample of size 100 was created by defining $y_1 = \hat{P}^{-1}(1/101)$; $y_2 = \hat{P}^{-1}(2/101)$; ...; $y_{100} = \hat{P}^{-1}(100/101)$. Then P15 was calculated from the empirical distribution function for the y's by the computer program listed in Andrews et al. By this program, the sample median m, a multiple of the sample median absolute deviation from the median d, and the 100 residuals from m were calculated. All pseudo-observations (bad), the residuals of which exceeded $1.5d$ in absolute value, were replaced by $1.5d$ times the sign of the residual; all other pseudo-observations (good) were left alone. The estimate was then the sum of the numbers so obtained divided by the total number of good pseudo-observations. If all 100 pseudo-observations were bad, the median was used. For purposes of this article, P15 was said to be incalculable if $p_1 > 1/101$ or $p_k < 100/101$.

2.2.7 Sine *Curve* M *Estimator* (AMT).

This estimator is also described in Andrews et al. (1972), and the computer program given there was used for calculations of this article. The calculations were on the same pseudosample as that created by P15. Let $\psi(y) = \sin(y/2.1)$ if $|y| \le 2.1\pi$; otherwise let it equal zero. Then the program solved the equation $\sum_1^{100} \psi((y_i - \hat{\mu})/d_1) = 0$ iteratively. A multiple of the median of the absolute deviations about the estimate $\hat{\mu}$ was used as the estimate of scale d_1. For purposes of this article AMT was said to be incalculable if $p_1 > 1/101$ or $p_k < 100/101$.

The estimators P15 and AMT were included in this study to explore the possibility that robust estimators of

location such as those described in Andrews et al. could be useful for estimating the ED50. Most robust estimators of location are functionals of an estimate of the population distribution function (Andrews et al. 1972, Ch. 3) and therefore could be expressed as functionals of $\hat{P}(x)$. Creating pseudosamples is not necessary, but this technique allowed the use of existing computer programs. In preliminary experimentation, I found that creating pseudosamples of size greater than 100 had no practical effect on analysis of the data discussed in Sections 3 and 4.

3. DEFICIENCIES AND SENSITIVITIES

3.1 Potential Deficiencies of the Estimators

The MLE estimator can behave erratically as is illustrated by the hypothetical data and associated estimates of Tables 1 and 2. The MLE is not calculable for data 1*A* but is calculable for data 1*B*. The occurrence of a single additional response made calculation of the MLE possible. This happened in spite of the fact that the single response at $i = 4$ could be viewed as an anomalous response. The data 1*A* better conform to the logistic model than do data 1*B*, but the logistic model MLE iterative procedure fails on data 1*A* and works on data 1*B*. Notice that the true MLE's for data 1*A* are $\hat{\mu} = 6$, $\hat{\beta}^{-1} = 0$, estimates that are intuitively unsatisfactory unless one is willing to believe that the variance of the tolerance distribution nearly equals zero.

1. Hypothetical Data Representing Outcomes of Three Replications of a Biological Assay Experiment

Group	i:	1	2	3	4	5	6	7	8	9	10
Dose	x_i:	1	2	3	4	5	6	7	8	9	10
No. of Individuals	n_i:	20	20	20	20	20	20	20	20	20	20
Data 1A	r_{Ai}:	0	0	0	0	0	1	20	20	20	20
Data 1B	r_{Bi}:	0	0	0	1	0	1	20	20	20	20
Data 1C	r_{Ci}:	0	0	0	1	0	3	20	20	20	20

A deficiency of the often-used MLCS estimator is that, in duplicate experiments, the experiment yielding more responses may have the larger ED50 estimate. This bad feature of MLCS does not seem to be discussed in the statistical literature. Consider the data and estimates of Tables 1 and 2. The three data sets could be replications of the same experiment. Notice that $r_{Ai} \leq r_{Ci}$ for every $i = 1, \ldots, k$. Because the response rates in 1*C* are higher than those of 1*A*, one would expect the ED50 estimate for 1*C* to be no larger than that for 1*A*. Surprisingly, the MLCS and MLCSM estimates are larger for data 1*C* than for data 1*A*. This behavior of the MLCS technique is created by the definition of the empirical logit transformation when the response proportions are near 0 or 1. I know of no version of the empirical logit transformation that avoids this deficiency.

As Table 2 demonstrates, AMT sometimes produces a larger ED50 estimate for the duplicate experiment having the greater mortality. The SK and trimmed SK estimators do not have this deficiency. If one judged the MLE to be calculable for data 1*A*, then the MLE procedure would possess this deficiency.

3.2 Sensitivity of the Estimators to an Anomalous Response

A subject having an especially low tolerance could possibly be randomly assigned to a low-dose group; that subject would respond although there are no other responding subjects in that dose group and perhaps there are no responses in the next-higher dose groups. An anomalous response could also be attributable to a data-recording error, the administration of a miscalibrated dose, or some other technical error. The statistical methodology should not be so delicate that it is seriously affected by an anomalous response.

In order to determine the sensitivity of ED50 estimators to a single anomalous response, the following calculations were performed for each estimator. For a given set of data $\{(x_i, n_i, r_i), i = 1, \ldots, k\}$ the estimate $\hat{\mu}$ was found. Then the data were augmented by adding 1 to r_j and n_j for a chosen group j. The estimate of μ, $\hat{\mu}^{(j)}$, was calculated and the reduction, $\hat{\mu} - \hat{\mu}^{(j)}$, attributable to the additional response, was recorded. These reductions were calculated for $j = 1, 2, \ldots, k$. Some typical results are presented in Table 3.

A negative reduction indicates that the estimator produced a larger ED50 estimate when a response was added to the original data. The estimators MLCS, MLCSM, and AMT sometimes produced negative reductions. These negative reductions occurred even when the added response was not inconsistent with the original data (see Table 3, Data Set 3A).

A robust estimator should yield a smaller reduction for an obviously anomalous additional response than for a plausible additional response. This desirable property is possessed by the trimmed SK and AMT estimators. As the dose group index j varied, however, from dose groups in which an additional response could plausibly occur to implausibly low-dose groups, the reductions associated with MLCS and MLCSM steadily increased and the

2. Estimates of μ for Data of Table 1

Method:	*MLE*	*MLCS*	*MLCSM*	*SK*	*5%SK*	*10%SK*	*20%SK*	*50%SK*	*P15*	*AMT*
Data 1A	n.c.*	6.41	6.41	6.45	6.47	6.47	6.47	6.47	6.45	6.47
Data 1B	6.39	6.95	6.90	6.40	6.47	6.47	6.47	6.47	6.41	6.48
Data 1C	6.29	6.82	6.77	6.30	6.37	6.40	6.41	6.41	6.31	6.41

* n.c. indicates that the iterative calculation procedure does not converge.

3. Reductions of the Original Estimate Attributable to a Single Additional Response at Dose x: $k = 10$, $x_i = i$, $n_i = 10$; Responses Listed Are $r_1, \ldots, r_{10}$

Responses	Data Set 3A 0,0,0,0,2,3,10,10,10,10					Data Set 3B 0,1,1,1,3,7,9,9,9,10				
Additional Response at	$x = 2$	$x = 3$	$x = 4$	$x = 5$	Original Estimate	$x = 2$	$x = 3$	$x = 4$	$x = 5$	Original Estimate
MLE	.10	.10	.10	.09	6.00	.10	.09	.08	.07	5.50
MLCS	−.62	−.30	−.07	.07	6.07	.14	.09	.05	.06	5.50
MLCSM	−.47	−.24	−.07	.07	6.07	.09	.07	.05	.06	5.50
SK	.10	.10	.09	.07	6.00	.09	.09	.08	.06	5.50
5%SK	.01	.02	.05	.08	6.05	.09	.09	.09	.07	5.50
10%SK	.01	.01	.03	.08	6.10	.09	.11	.10	.08	5.50
20%SK	0	0	0	.08	6.21	.01	.01	.03	.10	5.50
50%SK	0	0	0	0	6.29	0	0	0	.09	5.50
P15	.09	.08	.09	.07	6.00	.09	.08	.08	.06	5.50
AMT	−.06	−.09	−.15	.09	6.09	−.00	.05	.09	.10	5.50

reductions associated with MLE, SK, and P15 remained constant or increased slightly (see Table 3).

4. MONTE CARLO COMPARISONS OF ESTIMATIONS

4.1 Description of Simulation Experiments

Fifteen Monte Carlo simulation experiments have been conducted to compare the 10 estimators. In each simulation experiment $k = 10$ dose groups were used with n individuals per group and with equally spaced dose levels of $x_i = i$, $i = 1, 2, \ldots, 10$. Five different models for the tolerance distribution were used. For each of these five models, three experiments were conducted, one experiment for each of $n = 5, 10, 20$ subjects per dose group. For each of the 15 experiments, 1,800 sets of data were generated and each set of data was analyzed by all 10 estimation methods. The empirical MSE's of the estimators were then calculated and compared. Preliminary experimentation indicated that 1,800 data sets were required to have a good chance of detecting a 10 percent difference between MSE's.

The true probability of response, $P(x)$, is given in Table 4 for each of the five models and for each of the 10 dose levels. All five models are symmetric distributions around an ED50 of $\mu = 5.5$. Model I is a logistic tolerance distribution. Models II, III, and IV are contaminated logistic distributions in which a randomly chosen subject from the contaminated logistic population has probability .90 of coming from a logistic (mean = 5.5; standard deviation = σ) population and a probability of .10 of coming from a logistic (mean = 5.5; standard deviation = $c\sigma$) population. For Model II, $c = 3$; for Model III, $c = 5$; and for Model IV, $c = 10$. Models II, III, and IV are denoted in the tables by $.9L(\sigma) + .1L(c\sigma)$, $c = 3, 5$, or 10. Model V is a degenerate contaminated logistic distribution in which a randomly chosen subject has probability .80 of coming from a logistic (mean = 5.5) population, probability .10 of having a tolerance of 1.5 dose units, and probability of .10 of having a tolerance of of 9.5 dose units. The scale parameters for Models I through IV were chosen so that $P(3) = .01$ and $P(8) = .99$. The experiments simulated in this Monte Carlo study were intended to represent the kind of data routinely collected at the Fisheries Bioassay Laboratory, Montana State University (Hamilton, Russo, and Thurston 1977, 1978). Such bioassays usually involve 4 to 10 dose levels. Ten dose levels were used in this simulation study to give the force of contamination a chance to affect the estimators.

4. Values of P(x), the Probability of a Response at Dose x, for Each of Five Models

Dose	Model I	II	III	IV	V
x = 1	.00026	.00157	.00187	.00188	.00020
x = 2	.00160	.00389	.00440	.00441	.10128
x = 3	.01000	.01000	.01000	.01000	.10800
x = 4	.05969	.03415	.02232	.02111	.14775
x = 5	.28516	.21945	.12904	.05018	.32813
x = 6	.71484	.78055	.87096	.94984	.67187
x = 7	.94031	.96585	.97768	.97889	.85225
x = 8	.99000	.99000	.99000	.99000	.89200
x = 9	.99840	.99611	.99560	.99559	.89872
x = 10	.99974	.99843	.99813	.99812	.99980

To generate the number of responses, r, for given n and $P(x)$, a uniform (0, 1) pseudorandom variate was generated by the multiplicative congruential method (Naylor et al. 1966) and the corresponding value of r was found by reference to the cumulative distribution function associated with (2.1). These and all subsequent calculations were performed by the XDS Σ-7 computer at the Montana State University Computing Center. All programs were written in Xerox Data Systems Extended FORTRAN IV language.

4.2 Statistical Analysis of Simulated Experiments

Each of the 1,800 data sets in an experiment produced a 10-element vector of estimates $(\hat{\mu}_1, \ldots, \hat{\mu}_{10})'$, where $\hat{\mu}_i$ is the estimate of μ by method i, and the associated vector of squared errors $((\hat{\mu}_1 - 5.5)^2, \ldots, (\hat{\mu}_{10} - 5.5)^2)'$. Let $(M_1, \ldots, M_{10})'$ be the vector of true, unknown MSE's for the 10 estimators. Based on the 1,800 squared error vectors, simultaneous .05 level tests of the 45 null

5. Comparison of MSE's for Model I: Logistic

n = 5										
A. L = 1,514	SK	P15	MLE	5%SK	MLCSM	10%SK	MLCS	AMT	20%SK	50%SK
MSE(×100)	10.39	10.41	10.43	10.78	11.44	11.49	11.56	12.54	13.23	18.74
Relative MSE	1.00	1.00	1.00	1.04	1.10	1.11	1.11	1.21	1.27	1.80
B. No. Data Sets	1,797	1,797	1,517	1,798	1,800	1,800	1,800	1,797	1,800	1,800
MSE(×100)	10.77	10.78	10.41	11.03	11.82	11.57	11.94	12.50	12.99	17.58
Relative MSE	1.03	1.04	1.00	1.06	1.14	1.11	1.15	1.20	1.25	1.69
C. Results of .05 Level Simultaneous Tests										
n = 10										
A. L = 1,786	SK	P15	MLE	5%SK	MLCSM	10%SK	AMT	MLCS	20%SK	50%SK
MSE(×100)	5.57	5.58	5.61	5.80	6.04	6.20	6.23	6.36	6.79	10.20
Relative MSE	1.00	1.00	1.01	1.04	1.09	1.11	1.12	1.14	1.22	1.83
B. No. Data Sets	1,789	1,789	1,777	1,799	1,800	1,800	1,789	1,800	1,800	1,800
MSE(×100)	5.70	5.70	5.60	5.89	6.15	6.28	6.34	6.47	6.85	10.22
Relative MSE	1.02	1.02	1.00	1.05	1.10	1.12	1.13	1.16	1.22	1.83
C. Results of .05 Level Simultaneous Tests										
n = 20										
A. L = 1,783	SK	P15	MLE	5%SK	MLCSM	MLCS	AMT	10%SK	20%SK	50%SK
MSE(×100)	2.70	2.71	2.72	2.86	2.88	2.98	3.06	3.10	3.34	4.15
Relative MSE	1.00	1.00	1.01	1.06	1.06	1.10	1.13	1.15	1.23	1.53
B. No. Data Sets	1,784	1,784	1,799	1,800	1,800	1,800	1,784	1,800	1,800	1,800
MSE(×100)	2.71	2.72	2.72	2.87	2.88	2.99	3.06	3.11	3.34	4.15
Relative MSE	1.00	1.00	1.00	1.06	1.06	1.10	1.13	1.15	1.23	1.53
C. Results of .05 Level Simultaneous Tests										

NOTE: Description of entries:
A. MSE's calculated by using those L data sets for which no estimator was incalculable. Each relative MSE is the ratio of the corresponding MSE to the smallest MSE.
B. MSE calculations for an estimator use all data sets for which that estimator was calculable; the number of data sets is presented.
C. The simultaneous tests are based on the L data sets and MSE's of A. The solid line having an asterisk beneath estimator $\hat{\mu}$ underlines all estimators having MSE's not significantly different from the MSE for $\hat{\mu}$.

hypotheses, H_0: $M_i = M_\mu$, were conducted by using the Hotelling T^2 procedure (Miller 1966, p. 197, Eq. (36)).

The results of these analyses are presented in Tables 5 through 9. When testing for significant differences between MSE's, the computer used only those data sets in which all 10 estimators were calculable. The number of such data sets is denoted by L on each of the 15 presentations of Tables 5 through 9.

There are five rows of numerical summary statistics in each presentation. The first two rows pertain to those L data sets for which no method was incalculable. The middle row indicates, for each estimator, the number of data sets for which that estimator was calculable. In the last two rows, all data sets for which a particular estimator was calculable were used to calculate the MSE for that estimator. The relative MSE's were found by dividing the MSE for an estimator by the smallest MSE among the 10 estimators.

Each presentation of Tables 5 through 9 contains a graphical description of the results of simultaneous tests for significant differences between MSE's. These tests are to help the reader determine which differences among the ordered MSE's could be attributable to Monte Carlo sampling variation. To interpret the graphical presentations, find the horizontal line that contains an asterisk directly underneath the symbol for a selected estimator. This line underscores all estimators that have MSE's not significantly different from the MSE for the selected estimator.

Each of the models used in this study is such that the doses are symmetrically placed around μ. For these models and over the subset of the sample space in which all 10 estimates are calculable, the estimators should be unbiased. Thus, the comparisons of MSE's are essentially comparisons of variances. We know that bias attributable to asymmetric placement of doses produces a negligible contribution to the MSE of the conventional SK estimator when n is small (Finney 1964, p. 527). I believe that a similar negligible contribution of bias holds for the nonparametric estimators studied in this article. Thus, I

6. Comparison of MSE's for Model II: .9 Logistic (σ) + .1 Logistic (3σ)

$n = 5$										
A. L = 1,201	AMT	10%SK	5%SK	P15	SK	MLE	20%SK	MLCSM	MLCS	50%SK
MSE(×100)	7.39	7.49	7.50	7.84	7.88	7.92	7.94	8.96	9.14	9.88
Relative MSE	1.00	1.01	1.01	1.06	1.07	1.07	1.08	1.21	1.24	1.34
B. No. Data Sets	1,773	1,800	1,782	1,773	1,773	1,228	1,800	1,800	1,800	1,800
MSE(×100)	7.81	7.77	7.93	8.34	8.39	7.98	7.92	9.67	9.87	9.16
Relative MSE	1.01	1.00	1.02	1.07	1.08	1.03	1.02	1.24	1.27	1.18
C. Results of .05 Level Simultaneous Tests										
$n = 10$										
A. L = 1,664	20%SK	AMT	5%SK	10%SK	P15	SK	MLE	MLCSM	50%SK	MLCS
MSE(×100)	4.23	4.32	4.37	4.39	4.49	4.51	4.58	4.81	5.10	5.52
Relative MSE	1.00	1.02	1.04	1.04	1.06	1.07	1.08	1.14	1.21	1.31
B. No. Data Sets	1,800	1,760	1,799	1,800	1,760	1,760	1,704	1,800	1,800	1,800
MSE(×100)	4.30	4.49	4.51	4.49	4.70	4.72	4.56	5.08	5.11	5.84
Relative MSE	1.00	1.04	1.05	1.04	1.09	1.10	1.06	1.18	1.19	1.36
C. Results of .05 Level Simultaneous Tests										
$n = 20$										
A. L = 1,694	50%SK	20%SK	P15	SK	5%SK	AMT	10%SK	MLE	MLCSM	MLCS
MSE(×100)	1.81	1.90	2.14	2.15	2.15	2.17	2.17	2.19	2.45	2.76
Relative MSE	1.00	1.05	1.18	1.19	1.19	1.20	1.20	1.21	1.35	1.53
B. No. Data Sets	1,800	1,800	1,695	1,695	1,799	1,695	1,800	1,799	1,800	1,800
MSE(×100)	1.80	1.91	2.14	2.15	2.17	2.17	2.18	2.22	2.52	2.92
Relative MSE	1.00	1.06	1.19	1.19	1.20	1.20	1.21	1.24	1.40	1.62
C. Results of .05 Level Simultaneous Tests										

NOTE: Description of entries: (Same as Table 5).

believe that the MSE's of Tables 5 through 9 are close approximations to the MSE's that would occur if the doses were asymmetrically placed around μ, as long as the dose range is wide enough (to enclose the first percentile and the 99th percentile of the tolerance distribution, say) and the dose mesh is fine enough so that several dose levels are between the first to 99th percentile.

4.3 Comparison of Estimators

The MLE was the most frequently incalculable estimator, being incalculable for 4,430 of the 27,000 (=15 × 1,800) data sets. In 3 of the 15 experiments I examined all data sets for which MLE was incalculable. Without exception, these data sets were of the kinds described in the last paragraph of Section 2.2.1.

The next most frequently incalculable estimators were SK, P15, and AMT. These estimators were incalculable for 624 of the data sets. Typically, MLE was calculable more often than SK, P15, and AMT when $n = 20$ and calculable less often when $n = 5$, 10. The MLCS estimators were calculable for all data sets. The estimators 50%SK and 20%SK were also always calculable; 10%SK was incalculable for only five data sets, and 5%SK was incalculable for 76 of the data sets.

Remember that a conservative policy regarding calculability of the nonparametric estimates was followed. The usual practice in calculating the SK estimate is to fabricate some data if $p_1 > 0$ or $p_k < 1$ (Finney 1964). For example, if $p_k < 1$, one usually assumes that the next dose in the series, although untested, would yield a response proportion of one. That dose, x_{k+1}, and associated hypothetical response, $s_{k+1} = 1$, are used in the analysis. If this practice had been employed, the nonparametric estimates would have been calculable for all simulated data sets.

The empirical MSE's will now be used to compare the estimators. There is usually no practical difference between MLE and SK. It is well known that, for the kinds of experiments simulated here, the SK estimate is a close approximation to the MLE (Church and Cobb 1973). The estimator P15 is a simplified Huber M estimator of location. The Huber M estimators were derived by Huber (1964) from general expressions for MLE's. It is not suprising that P15, SK, and MLE make up a single cluster of similarly behaving estimators. These estima-

7. Comparison of MSE's for Model III: .9 Logistic (σ) + .1 Logistic (5σ)

	n = 5									
A.	L = 830	AMT	20%SK	50%SK	10%SK	5%SK	P15	SK	MLE	MLCSM	MLCS
	MSE(×100)	3.49	3.50	3.60	4.30	5.08	6.34	6.47	6.51	7.82	8.20
	Relative MSE	1.00	1.00	1.03	1.23	1.46	1.82	1.86	1.87	2.24	2.35
B.	No. Data Sets	1,764	1,800	1,800	1,800	1,784	1,764	1,764	866	1,800	1,800
	MSE(×100)	3.99	3.63	3.56	4.27	4.87	5.73	5.83	6.70	7.68	8.07
	Relative MSE	1.12	1.02	1.00	1.20	1.37	1.61	1.64	1.88	2.16	2.27
C.	Results of .05 Level Simultaneous Tests										
	n = 10										
A.	L = 1,402	50%SK	20%SK	AMT	10%SK	5%SK	P15	SK	MLE	MLCSM	MLCS
	MSE(×100)	1.54	1.66	2.08	2.18	2.58	3.14	3.19	3.29	3.67	5.37
	Relative MSE	1.00	1.08	1.35	1.42	1.68	2.04	2.08	2.14	2.39	3.43
B.	No. Data Sets	1,800	1,800	1,753	1,800	1,797	1,753	1,753	1,449	1,800	1,800
	MSE(×100)	1.58	1.69	2.19	2.19	2.59	3.19	3.25	3.29	4.05	5.90
	Relative MSE	1.00	1.08	1.39	1.39	1.65	2.02	2.06	2.09	2.57	3.75
C.	Results of .05 Level Simultaneous Tests										
	n = 20										
A.	L = 1,638	50%SK	20%SK	10%SK	AMT	5%SK	P15	SK	MLE	MLCSM	MLCS
	MSE(×100)	.57	.63	.94	.99	1.19	1.40	1.43	1.55	2.52	3.63
	Relative MSE	1.00	1.10	1.63	1.73	2.08	2.44	2.49	2.69	4.39	6.33
B.	No. Data Sets	1,800	1,800	1,800	1,684	1,799	1,684	1,684	1,754	1,800	1,800
	MSE(×100)	.58	.64	.95	1.00	1.22	1.42	1.45	1.57	2.64	3.94
	Relative MSE	1.00	1.00	1.63	1.71	2.08	2.43	2.48	2.69	4.52	6.74
C.	Results of .05 Level Simultaneous Tests										

NOTE: Description of entries: (Same as Table 5).

tors had the smallest MSE's for the light-tailed tolerance distributions (Models I and V) but had large MSE's for heavy-tailed distributions (Model IV). The estimators SK and P15 usually had significantly smaller MSE's than MLE. The empirical MSE's for P15 were overall slightly smaller than for SK.

An interesting discovery of the Monte Carlo calculations is that MLCSM is better than MLCS. Before calculating the MLCS estimator of the ED50, one should adjust the observed response proportions. Although the MSE's for MLCS were less than 5 percent larger than the MSE's for MLCSM in the logistic Model I experiments, MLCS had up to 85 percent larger MSE's than MLCSM in the heavy-tailed tolerance distribution model experiments (Model IV). The MSE's for MLCSM were statistically significantly smaller than the MSE's for MLCS in all experiments but three (Models I, II, V, when $n = 5$).

The MLCS estimators had larger MSE's than did the MLE estimator of ED50; MLCS and MLCSM were especially poor when the tolerance distribution was heavy tailed (Model IV).

The MSE's for the trimmed SK estimators varied in a systematic way. If the tolerance distribution was light tailed (e.g., logistic Model I) the MSE increased as the amount trimmed increased; if the tolerance distribution was heavy tailed (e.g., Model IV), the MSE increased as the amount trimmed decreased. Relative to the other estimators, the performances of the trimmed SK estimators improved as n increased.

When the tolerance distribution was light tailed (Models I and V), 10%SK had less than a 15 percent larger MSE than P15, SK, and MLE; 20SK had less than a 25 percent larger MSE. For all other models either 10%SK or 20%SK had a smaller MSE than P15, SK, and MLE. For the heavy-tailed distribution (Model IV), the MSE's for P15, SK, and MLE were from 2.0 to 7.0 times the MSE's for 10%SK and 20%SK.

The estimator AMT behaved much like 10%SK. The MSE of AMT was statistically significantly smaller than the MSE of 10%SK only when $n = 5$ and the tolerance distribution was heavy tailed (Models III and IV). The MSE of AMT was statistically significantly larger than the MSE of 10%SK for $n = 20$ with Model IV and for $n = 5$, 10, 20 with Model V.

8. Comparison of MSE's for Model IV: .9 Logistic (σ) + .1 Logistic (10σ)

n = 5										
A. L = 600	50%SK	20%SK	AMT	10%SK	5%SK	P15	SK	MLE	MLCSM	MLCS
MSE(×100)	.80	.91	1.02	1.82	3.05	5.33	5.56	5.60	9.47	10.42
Relative MSE	1.00	1.15	1.28	2.28	3.82	6.68	6.97	7.02	11.88	13.06
B. No. Data Sets	1,800	1,800	1,764	1,800	1,787	1,764	1,764	636	1,800	1,800
MSE(×100)	.94	1.02	1.31	1.50	2.08	3.10	3.20	5.74	6.51	7.18
Relative MSE	1.00	1.08	1.39	1.59	2.23	3.29	3.41	6.11	6.92	7.64
C. Results of .05 Level Simultaneous Tests										
n = 10										
A. L = 1,045	50%SK	20%SK	AMT	10%SK	5%SK	P15	SK	MLE	MLCSM	MLCS
MSE(×100)	.32	.34	.48	.58	1.08	2.08	2.19	2.29	4.45	8.22
Relative MSE	1.00	1.06	1.49	1.81	3.37	6.50	6.84	7.17	13.91	25.68
B. No. Data Sets	1,800	1,800	1,753	1,800	1,798	1,753	1,753	1,092	1,800	1,800
MSE(×100)	.37	.39	.57	.58	.95	1.70	1.79	2.50	4.29	8.06
Relative MSE	1.00	1.05	1.53	1.56	2.55	4.56	4.79	6.18	11.49	21.61
C. Results of .05 Level Simultaneous Tests										
n = 20										
A. L = 1,442	50%SK	20%SK	10%SK	AMT	5%SK	P15	SK	MLE	MLCSM	MLCS
MSE(×100)	.13	.13	.17	.20	.34	.77	.82	.94	5.60	9.13
Relative MSE	1.00	1.01	1.32	1.56	2.61	5.86	6.26	7.21	42.77	69.75
B. No. Data Sets	1,800	1,800	1,800	1,682	1,800	1,682	1,682	1,560	1,800	1,800
MSE(×100)	.14	.14	.19	.22	.34	.74	.79	.97	6.33	11.13
Relative MSE	1.00	1.01	1.29	1.53	2.42	5.21	5.58	6.80	44.58	78.37
C. Results of .05 Level Simultaneous Tests										

NOTE: Description of entries: (Same as Table 5).

4.4 Empirical vs. Theoretical Properties of SK

One can ascertain the extent of agreement between theoretical and empirical results for SK. The estimator SK was said to be calculable when and only when $s_1 = .0$ and $s_{10} = 1.0$. For the experiments simulated here, the expected number of data sets for which SK is calculable is given by $1{,}800[1 - P(1)]^n[P(10)]^n$. These expected numbers are shown in Table 10 along with the two-tailed p values for testing whether the observed proportions of data sets for which SK was calculable are different from the expected proportions.

For the simulated experiments, whenever $s_1 = .0$ and $s_{10} = 1.0$, SK can be written as $9.5 - \sum_{i=2}^{9} s_i$. The MSE of SK, conditional on being calculable, is $\sum_{i=2}^{9} P(i)[1 - P(i)]/n$. The numerical values of these theoretical MSE's are shown in Table 10 along with two-tailed p values for testing whether the observed MSE's are different from the theoretical.

The comparisons of Table 10 indicate that the Monte Carlo derived values of Tables 5 through 9 are reasonably accurate, except possibly for Models II, III, and IV when $n = 10$ where the MSE's are positively biased. I feel that the relative MSE's would be little altered if the experiments at $n = 10$ were repeated with new sets of simulated data.

5. RECOMMENDATIONS AND DISCUSSION

5.1 Recommendations

Consider an experiment in which there are approximately 10 equally spaced doses with n subjects per dose, the dose range encloses at least the 1.0 and 99.0 percentiles of the tolerance distribution, and some dosages lie between the 1.0 and 99.0 percentiles. The purpose of the experiment is to estimate the ED50.

Taking into account (a) the relative MSE's of the estimator for a variety of tolerance distribution models, (b) the sensitivity of the estimator to an anomalous response, and (c) the ability of the estimator to produce calculable estimates, the trimmed SK's, especially 10%SK and 20%SK, are the recommended methods for $n = 10, 20$.

If sensitivity of the estimator to an anomalous response is not an important consideration, then AMT is as good as the trimmed SK estimators when $n = 5$ or 10.

The nonparametric estimators P15 and SK are pre-

9. Comparison of MSE's for Model V: Degenerate Contaminated Logistic

n = 5										
A. L = 1,784	SK	P15	5%SK	10%SK	MLE	20%SK	MLCSM	MLCS	AMT	50%SK
MSE(×100)	21.46	21.51	22.17	22.58	22.93	23.39	24.56	25.47	27.62	29.96
Relative MSE	1.00	1.00	1.03	1.05	1.07	1.09	1.14	1.19	1.29	1.39
B. No. Data Sets	1,795	1,795	1,795	1,796	1,789	1,800	1,800	1,800	1,795	1,800
MSE(×100)	21.40	21.44	22.10	22.50	22.96	23.29	24.51	25.42	27.48	29.93
Relative MSE	1.00	1.00	1.03	1.05	1.07	1.09	1.15	1.19	1.28	1.40
C. Results of .05 Level Simultaneous Tests										
n = 10										
A. L = 1,791	MLCSM	SK	P15	MLE	5%SK	20%SK	10%SK	MLCS	AMT	50%SK
MSE(×100)	11.08	11.15	11.22	11.84	12.18	12.39	12.70	12.76	15.17	16.66
Relative MSE	1.00	1.01	1.01	1.07	1.10	1.12	1.15	1.15	1.37	1.50
B. No. Data Sets	1,800	1,791	1,791	1,800	1,793	1,800	1,799	1,800	1,791	1,800
MSE(×100)	11.10	11.15	11.22	11.87	12.18	12.37	12.71	12.77	15.17	16.58
Relative MSE	1.00	1.00	1.01	1.07	1.10	1.11	1.15	1.15	1.37	1.49
C. Results of .05 Level Simultaneous Tests										
n = 20										
A. L = 1,792	SK	P15	MLCSM	MLE	20%SK	MLCS	5%SK	10%SK	50%SK	AMT
MSE(×100)	5.37	5.42	5.46	5.68	5.70	5.94	6.02	6.08	7.49	7.71
Relative MSE	1.00	1.01	1.02	1.06	1.06	1.11	1.12	1.13	1.39	1.43
B. No. Data Sets	1,792	1,792	1,800	1,800	1,800	1,800	1,794	1,800	1,800	1,792
MSE(×100)	5.37	5.42	5.48	5.70	5.72	5.96	6.03	6.12	7.50	7.71
Relative MSE	1.00	1.01	1.02	1.06	1.07	1.11	1.12	1.14	1.40	1.44
C. Results of .05 Level Simultaneous Tests										

NOTE: Description of entries: (Same as Table 5).

ferable to MLE, although they are less often calculable than MLE when $n = 20$. The estimators P15, SK, and MLE are recommended only if the tolerance distribution is light tailed (e.g., logistic).

10. Theoretical Results for SK As Compared With Values in Tables 5 Through 9

	Theoretical Value			p Value		
	n					
Model	5	10	20	5	10	20
	Number of Data Sets for Which SK Is Calculable					
I	1,795.3	1,790.7	1,781.4	.63	.67	.64
II	1,771.9	1,744.3	1,690.4	.94	.032	.69
III	1,766.4	1,733.5	1,670.2	.69	.015	.22
IV	1,766.4	1,733.5	1,669.5	.72	.013	.27
V	1,796.4	1,792.8	1,785.7	.59	.59	.10
	Conditional MSE (× 100) Given SK Is Calculable					
I	10.86	5.43	2.72	.81	.19	.94
II	8.72	4.36	2.18	.21	.025	.67
III	5.94	2.97	1.49	.60	.012	.50
IV	3.29	1.65	.82	.46	.029	.25
V	21.35	10.67	5.34	.94	.20	.86

It is recommended that the MLCS calculations always be preceded by a monotone adjustment of the response proportions. The estimator MLCSM is not recommended if the tolerance distribution is heavy tailed.

5.2 Discussion

In this article some robust methods for estimating the location parameter have been adapted to the problem of estimating the ED50 from binary-response bioassay data. The adaptation amounts to expressing the estimator as a functional of an estimate of the tolerance distribution function rather than as a functional of an empirical distribution function. The trimmed SK estimator that is an adaption of the trimmed mean is shown to be useful. Many other robust estimators are available; most of them could be altered to estimate the ED50.

As for further work, it would be interesting to know the properties of the robust estimators in asymmetric situations, that is, where there is asymmetric contamination or where the doses are not symmetric around μ, perhaps with $P(x_1) = .2$ or .4. Specific algorithms for calculating the robust estimates from data should be described rather than resorting to the pseudosample technique.

One may possibly adapt the robust procedures to more general problems, such as when the response curve is bounded between the unknown constants γ and τ, where $0 < \gamma < \tau < 1$. I am presently studying properties of estimates of the standard error of the trimmed SK estimate and of associated confidence interval estimates. The confidence interval procedure in Hamilton, Russo, and Thurston (1977) seems to be of practical value.

[*Received December 1977. Revised September 1978.*]

REFERENCES

Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rodgers, W.H., and Tukey, J.W. (1972), *Robust Estimates of Location—Survey and Advances*, Princeton, N.J.: Princeton University Press.

Ashton, Winifred D. (1972), *The Logit Transformation*, New York: Hafner Publishing Co.

Barlow, R.E., Bartholomew, D.J., Bremner, J.M., and Brunk, H.D. (1972), *Statistical Inferences Under Order Restrictions*, New York: John Wiley & Sons.

Church, J.D., and Cobb, E. Benton (1973), "On the Equivalence of Spearman-Karber and Maximum Likelihood Estimates of the Mean," *Journal of the American Statistical Association*, 68, 201–202.

Committee on Methods for Toxicity Tests With Aquatic Organisms (1975), *Methods for Acute Toxicity Tests With Fish, Macroinvertebrates, and Amphibians*, Corvallis, Ore.: U.S. Environmental Protection Agency Ecological Research Series, EPA-660/3-75-009.

Cox, D.R. (1970), *The Analysis of Binary Data*, London: Methuen & Co.

Finney, D.J. (1964), *Statistical Method in Biological Assay, Second Ed.*, New York: Hafner Publishing Co.

Hamilton, Martin A., Russo, Rosemarie C., and Thurston, Robert V. (1977), "Trimmed Spearman-Karber Method for Estimating Median Lethal Concentrations in Toxicity Bioassays," *Environmental Science and Technology*, 11, 714–719; Correction (1978), 12, 417.

Huber, P.J. (1964), "Robust Estimation of a Location Parameter," *Annals of Mathematical Statistics*, 35, 73–101.

Miller, Rupert G., Jr. (1966), *Simultaneous Statistical Inference*, New York: McGraw-Hill Book Co.

Naylor, Thomas H., Balintfy, Joseph L., Burdick, Donald S., and Chu, Kong (1966), *Computer Simulation Techniques*, New York: John Wiley & Sons.

Stephan, Charles E. (1977), "Methods for Calculating an LC50," in *Aquatic Toxicology and Hazard Evaluation. ASTM STP 634*, eds. F.L. Mayer and J.L. Hamelink, Philadelphia: American Society for Testing and Materials, 65–84.